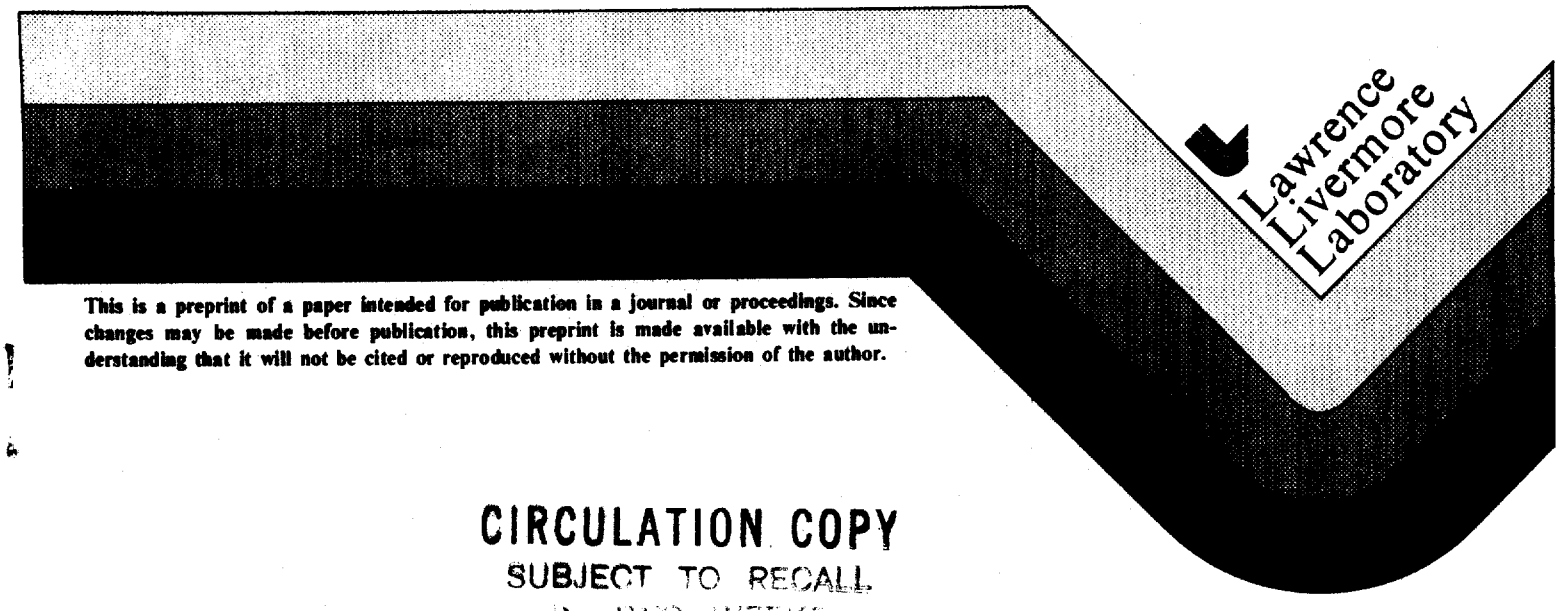
UCRL-84566
PREPRINT

# An Archival Memory System (AMS) Specification

George A. Michael
Lawrence Livermore National Laboratory
P.O. Box 808
Livermore, California

4th IEEE Workshop on Mass Storage
Denver, Colorado
March 16-17, 1980



# AN ARCHIVAL MEMORY SYSTEM (AMS) SPECIFICATION

G. A. Michael

Lawrence Livermore National Laboratory
Livermore, California

## ABSTRACT

The technical specification for an archival memory system is described. A phased acquisition leading ultimately to a very large (>$10^{15}$ bits) on line memory facility is envisioned.

## INTRODUCTION

The Lawrence Livermore National Laboratory (LLL) has been searching for a new archival memory system for over two years. The various steps taken have been reported elsewhere, Michael[1-3] but for the sake of continuity these are summarized here:

1. The needs of the large D.O.E. national laboratories were surveyed and a Request for Information (RFI) was sent to a large group of manufacturers. The goals were to quantify storage needs and to show that there was more of a market than many industry representatives had thought.

2. Based on known future requirements as well as product capabilities, a general system design was published. Reliability, longevity, modularity, and usability were the goals of the design.

There are two remaining steps to take:

3. We must develop a specification that is both precise and general for an AMS.

4. We must develop a complete software plan (External Reference Specification) to provide full storage management within a (possibly) distributed network of computer resources.

The third item listed above is the subject of this paper.

## THE SYSTEM SPECIFICATION

We are told that it is unusual to air the contents of a technical specification prior to its issuance. Among the reasons for doing so are these three:

First, an AMS will be one of the most important facilities in future computer systems. As such, it is useful to attempt to stimulate discussion of its characteristics by both users and manufacturers.

Second, if commonality of hardware, software and management procedures can be maximized, both users and manufacturers stand to benefit.

Third, if there are serious flaws in our plans, there is a better chance they will be discovered by publishing the plans before we lose our flexibility to adjust things.

The overall specifications for the Archival Memory System (AMS) are given in figures 1 to 8. The rest of this presentation consists of comments to amplify some of the material listed in these figures.

Figure 1

- The plan to acquire an AMS should be seen as (possibly) being carried out over more than one year. We concede that the timetable given may be optimistic.

- By archival we mean that the data must be readable on the order of 20 years.

- We are hoping that the equipment we acquire will be commercially available -- not custom made for us.

- Certain terms are defined:

  Volume: A volume is a unit of mediam on which data are written or recorded and from which data once written are read. A volume may contain an arbitrary number of "files" (see definition of File).

  On-line: A volume is said to be on-line if it can be accessed without human intervention.

  Off-line: Some sort of human operator action is required, e.g., a unit of the

mediam which as been removed from the on-line system must be manually re-inserted.

On-line Library: A facility used for the on-line storage of volumes not directly mounted on read/write or read-only Stations from which volumes can be automatically retrieved and automatically mounted upon a Station.

Shelf Library: Literally a suitable place for storing those Volumes no longer in general use. Human intervention is required to retrieve any such Volumes and re-insert them into the On-Line Library.

User Data: The data bits directly useful to the user, exclusive of identification bits, error detection and correction bits, and housekeeping bits.

File: A collection of user data referred to by a unique name. The actual "file" recorded in the AMMS may include headers and identification data to assure data privacy and security that also could be used to reconstruct file directories if need be, and trailing checksums or the like to increase data integrity.

Archival: An archival system is a write-once-read thereafter system with the highest degree of certainty for correct data recovery.

Postable: A volume is said to be postable if a major interruption may occur in the recording of data without damage to previously recorded data. During the interruption it shall be possible to perform any operation that may be performed with a fully recorded volume such as dismounting and re-mounting in any recording mechanism.

Read-Only Station: A Station capable of performing read functions, but not capable of performing write functions.

Partitioned: If the system is so divided or compartmentalized that some on-line volumes are not accessible by all Stations, it is said to be partitioned.

Figure 2

- The environment for the AMS is shown in outline form. It is to be seen as a part of a generalized storage facility (GSF) most of which should be centralized.
- If the need arises, portions of the storage facility can be moved to remote sites.
- Many of the network issues involving security, access and confidentiality directly affect the GSF.
- The only process not serviced by a machine of one type or another is that of fetching volumes from a shelf for insertion into the on-line store.

Figure 3

- The purpose of calling out separate controllers and interfaces is to allow as many simultaneous accesses as possible. In particular, non-data transmit operations should not tie up the data channels.
- At least for LLL a file management system already exists. It will be expanded and the AMS will be integrated into it. However, a vendor is permitted to offer his own system.
- Vendors are expected to provide all the low level utilities such as hardware drivers and preventative maintenance routines. Such routines must be capable of being invoked from an operations and control console. Such consoles must generally be usable without pulling the whole system off-line.

Figure 4

- The first increment of the AMS ($\geq 1.5 \times 10$" bits) is to be sought during fiscal year 1981.
- Some constraints on Volume size are given.
- A volume being postable means that its full capacity need not be written all at once. Instead it must be capable of having data added on demand and of being moved and remounted. Ideally it must function correctly on any read/write or read only station.
- Any environmental constraints must be stated.
- Those on-line volumes that are not directly mounted for computer access but can be obtained without human intervention are said to be in a "Juke Box" serviced by a "Robot".
- We allow the possibility that each Robot can get to only a subset of the totality of on-line volumes. However, all volumes must be accessible by at least one Robot. Such a configuration is said to be partitioned.

Figure 5

- Stations are of two types: those having both write and read capabilities and those able to read only. By implication a read only station must not have any ability to accidentally write on a volume. The LLL strategy is to never place a volume in a read/write station once it has been completely filled. This is obviously to

reduce the danger of loss of data due to inadvertent writing.

- If it is possible, the read and the write capabilities should be separate and independent features in a read/write station.
- Volumes must be usable among all stations of both the read/write and read only classes.
- The stations should be capable of being accessed over more than one channel.

Figure 6

- The time and performance limits given are to be taken as "worst case only". We certainly hope for much better performance.
- The same is true for the reliability range.

Figure 7

- There are three general comments about errors: First, one would hope that actual system performance will be orders of magnitude better than the worst case limits given. Second, while errors will occur and be corrected by retrying, we want all errors to be logged and the record thereof presented on request. Third, while truly heroic mechanisms should be designed and implemented to avoid undetected errors, no errors should be allowed to pull the whole system off line.
- We will request complete error characteristics for all AMS components.
- The desired environment for the AMS is a computer room but if special ranges are needed, they may be acceptable.
- There are three parts to the acceptance test. The first is in the manufacturer's facilities, the second is after delivery and installation, and the third requires a contiguous 30-day performance that meets all the requirements described in the specification. The contiguous 30-day test must occur within a 90-day window.

Figure 8

- After vendor selection and prior to delivery, the full ground rules for the contiguous 30-day test will be negotiated with the vendor. Shown in this figure are some of the things considered.

## CONCLUSION

It is our desire to solicit for an AMS during the next fiscal year (FY81). At present unfortunately it is not possible to give a precise date for starting. The contents of the specification given in this paper constitute a best estimate as they are now seen. Changes are possible but major ones do not seem probable.

## ACKNOWLEDGMENT

Archival Mass Storage System has been a reality at LLL for over ten years. It is the result of the work of many people: G. Boer, S. Coleman, J. Fletcher, B. Young, and M. Zosel to name some. S. Stone translated our current ideas about the AMS into the specifications we have been discussing. One of the most basic components in the existing LLL configuration is the IBM 1360, the Photostore. It is entirely appropriate to acknowledge the skills of J. Dimmick of IBM (now retired) whose skills made our Photostore by all reports, the best installation of its kind in the country.

## REFERENCES

1. Proceedings, D.O.E./NCAR Mass Storage Workshop, Dec. 8-9, 1977, Boulder, Colorado, A. V. Kerr, B. T. O'Lear, Chairman.
2. G. A. Michael, "Mass Archival Storage: Some Trends, Needs and Plans at D.O.E. Laboratories", JAPE, Vol. 6, No. 1, February 1980, pp 19-26.
3. G. A. Michael, "Archival Storage Requirements", Proceedings SPIE, Optical Information Storage, Vol. 177, pg 122-127.

"Work performed under the auspices of the U.S. Department of Energy by the Lawrence Livermore Laboratory under contract number W-7405-ENG-48."

SPECIFICATION FOR AN

ARCHIVAL MASS MEMORY SYSTEM

PURPOSE: INITIAL PHASE OF A

MULTI-STEP PROCUREMENT OF AN

- ON LINE ARCHIVAL MASS MEMORY SYSTEM.

TYPE: (PROBABLY) OPTICAL, DIGITAL, - [OTHERS POSSIBLE]

COMMERCIALLY AVAILABLE

MODULAR, AUTOMATIC

AVAILABLE: 24 HRS/DAY 365 DAYS/YR

ARCHIVAL => ≥20 YRS

DEFINITIONS: VOLUME, ON-LINE/OFF-LINE

USER DATA/FILE/STATIONS, ON-LINE LIBRARY

SHELF LIBRARY

CAPACITY

| 1981 - 82 | 1982 - 83 | 1983 - 85 |
|---|---|---|
| $>10^{11}$ | $>10^{13}$ | $>10^{15}$ |

DETERMINATION OF LONG TERM AVAILABILITY

OF ALL COMPONENTS

FIGURE 1

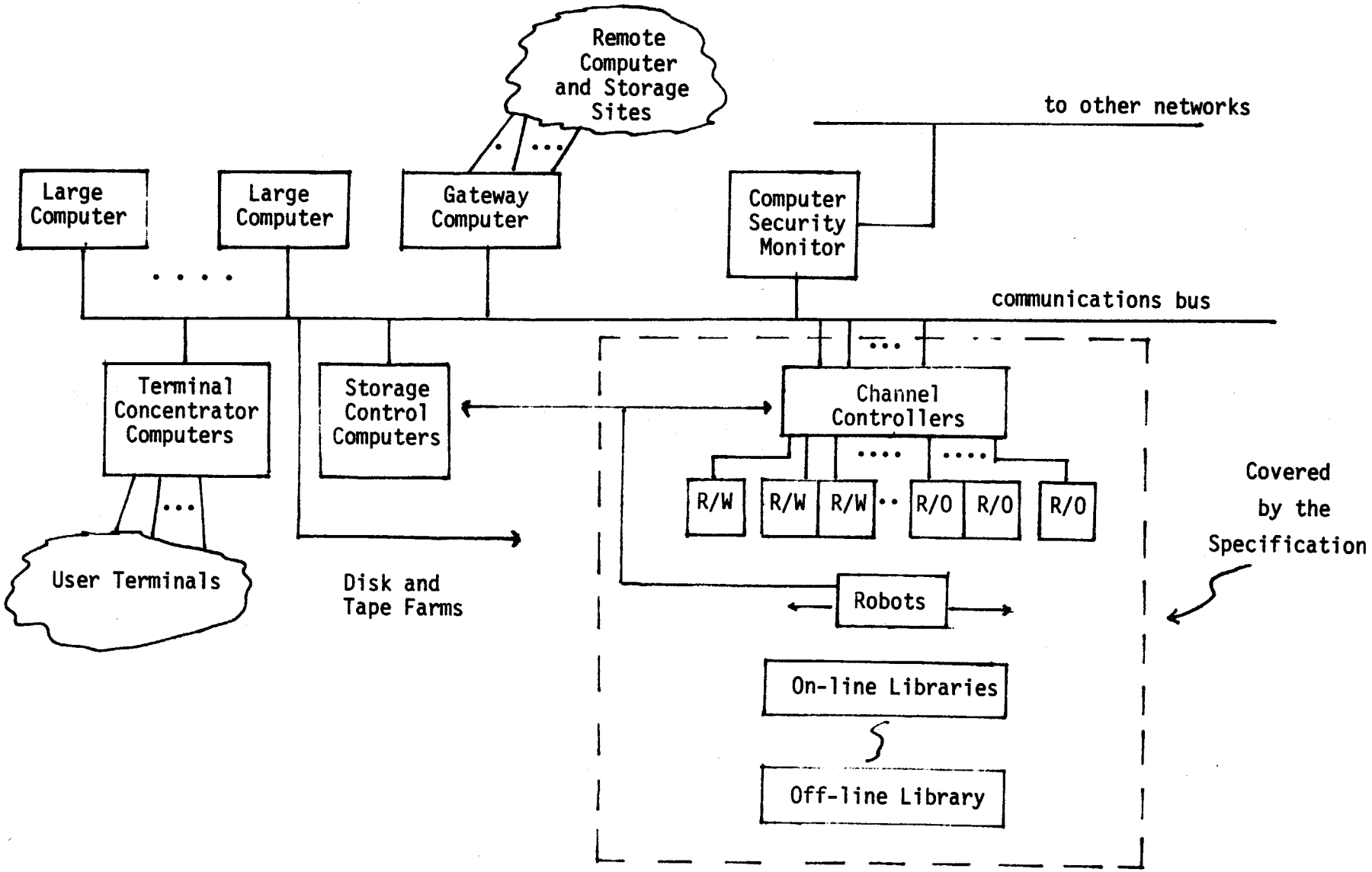

Remote Computer and Storage Sites
to other networks
Large Computer
Large Computer
Gateway Computer
Computer Security Monitor
communications bus
Terminal Concentrator Computers
Storage Control Computers
Channel Controllers
R/W
R/W
R/W
R/O
R/O
R/O
Covered by the Specification
User Terminals
Disk and Tape Farms
Robots
On-line Libraries
Off-line Library

Figure 2

## CONTROLLERS

ONE PER INTERFACE CHANNEL DESIRED

COMPUTER BASED

## INTERFACES

TWO INDEPENDENT CHANNELS REQUIRED IN '81

ABLE TO ATTACH TO OCTOPUS SYSTEM BUS

## SOFTWARE

EXISTING LLL FILE MANAGEMENT SYSTEM

OPTION OF VENDOR-SUPPLIED SYSTEM OR

UTILITIES, PREVENTIVE MAINTENANCE (ON-LINE)

SYSTEM DIAGNOSTIC CAPABILITIES VIA

OPERATIONS AND CONTROL STATIONS

FIGURE 3

1981FY

STORAGE CAPACITY > $1.5 \times 10^{11}$ BITS

VOLUME CAPACITY $10^9 < VC < 10^{11}$ BITS
(MULTIPLE SOURCES OF SUPPLY DESIRED)
LIFE TIME > 20 YRS ($10^5$ LOAD PASSES)
POSTABLE
UNIQUE IDENTIFIERS/HUMAN READABLE
ENVIRONMENTAL CONSTRAINTS

ON-LINE LIBRARY (JUKE BOX)
ACCESS MECHANISMS
PARTITIONING

FIGURE 4

STATIONS

MINIMUM OF THREE STATIONS IN '81 (AT LEAST ONE WRITE)

WRITING/READING ("D.R.A.W." TYPE FEATURE)

(EXTRA, INDEPENDENT READ CAPABILITY DESIRED)

READ ONLY (PLAYER)

NO ABILITY TO ACCIDENTALLY WRITE

ALL STATIONS EQUAL (ABLE TO READ ANY VOLUME)

SPECIAL DATA INTEGRITY (CHECKING)

MULTI-ACCESS FEATURE

FIGURE 5

## Access Provisions and Times

If applicable

Mount Volume from on-line library to designated station $\leq$ 5 sec. to full ready

Random read access to file on an already mounted volume < 1 sec.

Write access to file on an already mounted volume < 2 sec.

## Sustained Data Transfer Throughput Rate (user bits) Over a Volume

Across LLL-archive Mem interface > $5 \times 10^6$ b/sec.

Max peak rate < $4 \times 10^7$ b/sec.

No impact on rate from non-data transmit actions

Reliability

Available > 90%

Soft Degradation

Duplicate pathing

Figure 6

## Error Rates

All (recent) states reportable on request

Recoverable error < 1 in $10^4$ operations

Unrecoverable error < 1 in $10^{11}$ bits

report to nearest block + bad data if requested

(context recoverability)

Component errors recoverable or

unrecoverable ( = hard error)

## Environmental Range

Computer Room Conditions desired

Temp: 15°c to 30°c

Humidity (20 to 70)% relative

## Tests and Acceptance

In-plant

On delivery and installation

30-day acceptance (Contiguous) within 90 days

Figure 7

## The 30 Day Acceptance Test: A Negotiated Document

A. Describe the ground rules for conducting the test, the equipment configuration and the performance levels expected.

B. Definitions of terms:
   Effectiveness Level, Available Time, Down Time
   Preventive Maintenance Time, Uncounted Time.

C. Software: Manufacturer-Supplied; Laboratory-Supplied
   Diagnostics, Determination of problem responsibilities.

D. File size and frequency distributions.

E. Error classes, analysis, unanticipated condition handling.

F. Daily Reports and Performance Logging.

G. Media Certification.

H. Management of Equipment Updates (FCO's) and Replacement.

Figure 8